

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: RAMESCH V. NAIR
MARK S. PAYNE
DONALD E. TRIMBUR
FERNANDO VALLE

(ii) TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
GLYCEROL BY RECOMBINANT
ORGANISMS

(iii) NUMBER OF SEQUENCES: 43

(iv) CORRESPONDENCE ADDRESS:

(A) NAME: E. I. DU PONT DE NEMOURS AND COMPANY
(B) STREET: 1007 MARKET STREET
(C) CITY: WILMINGTON
(D) STATE: DELAWARE
(E) COUNTRY: U.S.A.
(F) ZIP: 19898

(A) ADDRESSEE: GENENCOR INTERNATIONAL, INC.
(B) STREET: 4 CAMBRIDGE PLACE
1870 SOUTH WINTON ROAD
(C) CITY: ROCHESTER
(D) STATE: NEW YORK
(E) COUNTRY: U.S.A.
(F) ZIP: 14618

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: DISKETTE, 3.5 INCH
(B) COMPUTER: IBM PC COMPATIBLE
(C) OPERATING SYSTEM: MICROSOFT WINDOWS 95
(D) SOFTWARE: MICROSOFT WORD VERSION 7.0A

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 60/030602
(B) FILING DATE: NOVEMBER 13, 1996
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: FLOYD, LINDA AXAMETHY
(B) REGISTRATION NUMBER: 33,692
(C) REFERENCE/DOCKET NUMBER: CR-9981-B

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 302-892-8112
(B) TELEFAX: 302-773-0164

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1380 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

| | | | | | | |
|------------|-------------|------------|------------|------------|------------|------|
| CTTTAATTTT | CTTTTATCTT | ACTCTCCTAC | ATAAGACATC | AAGAAACAAT | TGTATATTGT | 60 |
| ACACCCCCCC | CCTCCACAAA | CACAAATATT | GATAATATAA | AGATGTCTGC | TGCTGCTGAT | 120 |
| AGATTAAACT | TAACTTCCGG | CCACTTGAAT | GCTGGTAGAA | AGAGAAGTTC | CTCTTCTGTT | 180 |
| TCTTTGAAGG | CTGCCGAAAA | GCCTTCAAG | GTTACTGTGA | TTGGATCTGG | TAACTGGGGT | 240 |
| ACTACTATTG | CCAAGGTGGT | TGCCGAAAAT | TGTAAGGGAT | ACCCAGAAGT | TTTCGCTCCA | 300 |
| ATAGTACAAA | TGTGGGTGTT | CGAAGAAGAG | ATCAATGGTG | AAAAATTGAC | TGAAATCATA | 360 |
| AATACTAGAC | ATCAAAACGT | GAAATACTTG | CCTGGCATCA | CTCTACCCGA | CAATTGGTT | 420 |
| GCTAATCCAG | ACTGATTGA | TTCAGTCAAG | GATGTCGACA | TCATCGTTT | CAACATTCCA | 480 |
| CATCAATTAA | TGCCCGTAT | CTGTAGCAA | TTGAAAGGTC | ATGTTGATTC | ACACGTCAGA | 540 |
| GCTATCTCCT | GTCTAAAGGG | TTTGAAAGTT | GGTGCTAAAG | GTGTCCAATT | GCTATCCTCT | 600 |
| TACATCACTG | AGGAACCTAGG | TATTCAATGT | GGTGCTCTAT | CTGGTGCTAA | CATTGCCACC | 660 |
| GAAGTCGCTC | AAGAACACTG | GTCTGAAACA | ACAGTTGCTT | ACCACATTCC | AAAGGATTTC | 720 |
| AGAGGCGAGG | GCAAGGACGT | CGACCATAAG | GTTCTAAAGG | CCTTGTCCA | CAGACCTTAC | 780 |
| TTCCACGTTA | GTGTCATCGA | AGATGTTGCT | GGTATCTCCA | TCTGTGGTGC | TTTGAAGAAC | 840 |
| GTTGTTGCCT | TAGGTTGTGG | TTTCGTCGAA | GGTCTAGGCT | GGGTTAACAA | CGCTTCTGCT | 900 |
| GCCATCCAAA | GAGTCGGTTT | GGGTGAGATC | ATCAGATTG | GTCAAATGTT | TTTCCCAGAA | 960 |
| TCTAGAGAAC | AAACATACTA | CCAAGAGTCT | GCTGGTGTG | CTGATTTGAT | CACCACCTGC | 1020 |
| GCTGGTGGTA | GAAACGTCAA | GGTGCTAGG | CTAATGGCTA | CTTCTGGTAA | GGACGCCTGG | 1080 |
| GAATGTGAAA | AGGAGTTGTT | GAATGCCAA | TCCGCTCAAG | GTAAATTAC | CTGCAAAGAA | 1140 |
| GTTCACGAAT | GGTTGGAAAC | ATGTGGCTCT | GTCGAAGACT | TCCCATTATT | TGAAGCCGTA | 1200 |
| TACCAAATCG | TTTACAACAA | CTACCCAATG | AAGAACCTGC | CGGACATGAT | TGAAGAATTA | 1260 |
| GATCTACATG | AAGATTAGAT | TTATTGGAGA | AAGATAACAT | ATCATACTTC | CCCCACTTT | 1320 |
| TTCGAGGCTC | TTCTATATCA | TATTCTAA | TTAGCATTAT | GTCATTCTC | ATAACTACTT | 1380 |

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2946 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

| | | | | | | |
|-------------|-------------|--------------|------------|-------------|-------------|------|
| GAATTCTGAGC | CTGAAGTGCT | GATTACCTTC | AGGTAGACTT | CATCTTGACC | CATCAACCCC | 60 |
| AGCGTCAATC | CTGCAAATAC | ACCACCCAGC | AGCACTAGGA | TGATAGAGAT | AATATAAGTAC | 120 |
| GTGGTAACGC | TTGCCTCATC | ACCTACGCTA | TGGCCGGAAT | CGGCAACATC | CCTAGAATTG | 180 |
| AGTACGTGTG | ATCCGGATAA | CAACGGCACT | GAATATATCT | TCGGTATCGT | AAAGATGTGA | 240 |
| TATAAGATGA | TGTATACCCA | ATGAGGGAGCG | CCTGATCGT | ACCTAGACCT | TAGTGGCAAA | 300 |
| AACGACATAT | CTATTATAGT | GGGGAGAGTT | TCGTGCAAAT | AACAGACGCA | GCAGCAAGTA | 360 |
| ACTGTGACGA | TATCAACTCT | TTTTTTATTA | TGTAATAAGC | AAACAAGCAC | GAATGGGGAA | 420 |
| AGCCTATGTG | CAATCACCAA | GGTCGCCCT | TTTTTCCCAT | TTGCTAATTT | AGAATTAAA | 480 |
| GAAACCAAAA | GAATGAAGAA | AGAAAAACAAA | TACTAGCCCT | AACCCTGACT | TCGTTTCTAT | 540 |
| GATAATAACCC | TGCTTTAATG | AACGGTATGC | CCTAGGGTAT | ATCTCACTCT | GTACGTTACA | 600 |
| AACTCCGGTT | ATTTTATCGG | AACATCCGAG | CACCCGCGCC | TTCCCTCAACC | CAGGCACCGC | 660 |
| CCCAGGTAAC | CGTGCACGAT | GAGCTAATCC | TGAGCCATCA | CCCACCCCCAC | CCGTTGATGA | 720 |
| CAGCAATTCTG | GGAGGGCGAA | AATAAAACTG | GAGCAAGGAA | TTACCATCAC | CGTCACCATC | 780 |
| ACCATCATAT | CGCCTTAGCC | TCTAGCCATA | GCCATCATGC | AAGCGTGTAT | CTTCTAAAGAT | 840 |
| TCAGTCATCA | TCATTACCGA | GTGGTTTTTC | CTTCACATGA | TGAAGAAGGT | TTGAGTATGC | 900 |
| TCGAAACAAT | AAGACGACGA | TGGCTCTGCC | ATTGGTTATA | TTACGCTTTT | GC GGCGAGGT | 960 |
| GCCGATGGGT | TGCTGAGGGG | AAGAGTGTAT | AGCTTACGGA | CCTATTGCCA | TTGTTATTCC | 1020 |
| GATTAATCTA | TTGTTCAGCA | GCTCTTCTCT | ACCCCTGTAT | TCTAGTATTT | TTTTTTTTTT | 1080 |
| TTTTTGGTTT | TACTTTTTT | TCTTCTTGCC | TTTTTTCTT | GTTACTTTT | TTCTAGTTT | 1140 |
| TTTTCTTCC | ACTAAGCTTT | TTCCCTTGATT | TATCCTTGGG | TTCTTCTTTC | TACTCCTTTA | 1200 |
| GATTTTTTTT | TTATATATTA | ATTTTAAGT | TTATGTATTT | TGGTAGATTC | AATTCTCTTT | 1260 |
| CCCTTCCCTT | TTCCCTTCGCT | CCCCCTTCCCTT | ATCAATGCTT | GCTGTCAGAA | GATTAACAAG | 1320 |
| ATACACATTC | CTTAAGCGAA | CGCATCCGGT | GTTATATACT | CGTCGTGCAT | ATAAAATTTT | 1380 |

GCCTTCAAGA TCTACTTCC TAAGAAGATC ATTATTACAA ACACAAC TG ACTCAAAGAT 1440
 GACTGCTCAT ACTAATATCA AACAGCACAA ACACGTGTCAT GAGGACCATC CTATCAGAAG 1500
 ATCGGACTCT GCCGTGTCAA TTGTACATTT GAAACGTGCG CCCTTCAAGG TTACAGTGAT 1560
 TGGTTCTGGT AACTGGGGGA CCACCATCGC CAAAGTCATT GCGGAAAACA CAGAATTGCA 1620
 TTCCCATATC TTCGAGGCCAG AGGTGAGAAT GTGGGTTTT GATGAAAAGA TCGGCGACGA 1680
 AAATCTGACG GATATCATAA ATACAAGACA CCAGAACGTT AAATATCTAC CCAATATTGA 1740
 CCTGCCCAT AATCTAGTGG CCGATCCTGA TCTTTACAC TCCATCAAGG GTGCTGACAT 1800
 CCTTGTTC AACATCCCTC ATCAATTTC ACCAACATA GTCAAACAAAT TGCAAGGCCA 1860
 CGTGGCCCCCT CATGTAAGGG CCATCTCGTG TCTAAAAGGG TTCGAGTTGG GCTCCAAGGG 1920
 TGTGCAATTG CTATCCTCCT ATGTTACTGA TGAGTTAGGA ATCCAATGTG GCGCACTATC 1980
 TGGTGCAAAC TTGGCACCGG AAGTGGCCAA GGAGCATTGG TCCGAAACCA CCGTGGCTTA 2040
 CCAACTACCA AAGGATTATC AAGGTGATGG CAAGGATGTA GATCATAAGA TTTTGAAATT 2100
 GCTGTTCCAC AGACCTTACT TCCACGTCAA TGTCATCGAT GATGTTGCTG GTATATCCAT 2160
 TGCCGGTGCC TTGAAGAACG TCGTGGCACT TGCATGTGGT TTCGTAGAAG GTATGGATG 2220
 GGGTAACAAT GCCTCCGCAG CCATTCAAAG GCTGGGTTA GGTGAAATTAA TCAAGTCGG 2280
 TAGAATGTTT TTCCCAGAAT CCAAAGTCGA GACCTACTAT CAAGAATCCG CTGGTGTG 2340
 AGATCTGATC ACCACCTGCT CAGGGGTAG AAACGTCAAG GTGCCACAT ACATGGCCAA 2400
 GACCGGTAAG TCAGCCTGG AAGCAGAAAA GGAATTGCTT AACGGTCAAT CCGCCCAAGG 2460
 GATAATCACA TGCAGAGAAG TTCACGAGTG GCTACAAACA TGTGAGTTGA CCCAAGAATT 2520
 CCCAATTATT CGAGGCAGTC TACCAAGATAG TCTACAACAA CGTCCGCATG GAAGACCTAC 2580
 CGGAGATGAT TGAAGAGCTA GACATCGATG ACGAATAGAC ACTCTCCCC CCCCTCCCC 2640
 TCTGATCTT CCTGTTGCCT CTTTTCCCC CAACCAATT ATCATTATAC ACAAGTTCTA 2700
 CAAACTACTAC TAGTAACATT ACTACAGTTA TTATAATTCTT CTATTCTCTT TTTCTTTAAG 2760
 AATCTATCAT TAACGTTAAT TTCTATATAT ACATAACTAC CATTATACAC GCTATTATCG 2820
 TTTACATATC ACATCACCGT TAATGAAAGA TACGACACCC TGTACACTAA CACAATTAAA 2880
 TAATGCCCAT AACCTTTCT GTTATCTATA GCCCTTAAAG CTGTTCTTC GAGCTTTCA 2940
 CTGCAG 2946

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3178 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTGCAGAACT TCGTCTGCTC TGTGCCCATC CTCGCGGTTA GAAAGAACGCT GAATTGTTTC 60
ATGCGCAAGG GCATCAGCGA GTGACCAATA ATCACTGCAC TAATTCCCTT TTGCAACAC 120
ATACTTATAT ACAGCACCAAG ACCTTATGTC TTTTCTCTGC TCCGATACGT TATCCCACCC 180
AACTTTATT TCAGTTTGG CAGGGGAAAT TTCACAACCC CGCACGCTAA AAATCGTATT 240
TAAACCTAAA AGAGAACAGC CACAAATAGG GAACTTGGT CTAAACGAAG GACTCTCCCT 300
CCCTTATCTT GACCGTGCTA TTGCCATCAC TGCTACAAGA CTAAATACGT ACTAATATAT 360
GTTTCGGTA ACGAGAACAGA GAGCTGCCGG TGCAGCTGCT GCCATGGCCA CAGCCACGGG 420
GACGCTGTAC TGGATGACTA GCCAAGGTGA TAGGCCGTTA GTGCACAATG ACCCGAGCTA 480
CATGGTGCAA TTCCCCACCG CCGCTCCACC GGCAGGTCTC TAGACGAGAC CTGCTGGACC 540
GTCTGGACAA GACGCATCAA TTGACGTGT TGATCATCGG TGGCGGGGCC ACGGGGACAG 600
GATGTGCCCT AGATGCTGCG ACCAGGGGAC TCAATGTGGC CCTTGTGAA AAGGGGGATT 660
TTGCCTCGGG AACGTCGTCC AAATCTACCA AGATGATTCA CGGTGGGGTG CGGTACTTAG 720
AGAAGGCCTT CTGGGAGTTC TCCAAGGCAC AACTGGATCT GGTACATCGAG GCACTCAACG 780
AGCGTAAACA TCTTATCAAC ACTGCCCTC ACCTGTGCAC GGTGCTACCA ATTCTGATCC 840
CCATCTACAG CACCTGGCAG GTCCCGTACA TCTATATGGG CTGTAAATTCT TACGATTCT 900
TTGGCGGTTTC CAAAAACTTG AAAAATCAT ACCTACTGTC CAAATCCGCC ACCGTGGAGA 960
AGGCTCCCAT GCTTACCAACA GACAATTAA AGGCCTCGCT TGTGTACCAT GATGGGTCT 1020
TTAACGACTC GCGTTGAAC GCCACTTAG CCATCACGGG TGTGGAGAAC GGCGCTACCG 1080
TCTTGATCTA TGTCGAGGTA CAAAAATTGA TCAAAGACCC AACTCTGGT AAGGTTATCG 1140
GTGCCGAGGC CCGGGACGTT GAGACTAATG AGCTTGTCAAG AATCAACGCT AAATGTGTGG 1200
TCAATGCCAC GGGCCCATAAC AGTGACGCCA TTTGCAAAT GGACCGCAAC CCATCCGGTC 1260
TGCCGGACTC CCCGCTAAAC GACAACCTCA AGATCAAGTC GACTTCAAT CAAATCTCCG 1320
TCATGGACCC GAAAATGGTC ATCCCACCTA TTGGCGTTCA CATCGTATTG CCCTCTTTT 1380
ACTCCCCGAA GGATATGGGT TTGTTGGACG TCAGAACCTC TGATGGCAGA GTGATGTTCT 1440
TTTACCTTG GCAGGGCAAA GTCCTGCCG GCACCACAGA CATCCCACTA AAGCAAGTCC 1500

| | |
|---|------|
| CAGAAAACCC TATGCCTACA GAGGCTGATA TTCAAGATAT CTTGAAAGAA CTACAGCACT | 1560 |
| ATATCGAATT CCCCGTGAAG AGAGAAGACG TGCTAAGTGC ATGGGCTGGT GTCAGACCTT | 1620 |
| TGGTCAGAGA TCCACGTACA ATCCCCGAG ACGGGAAGAA GGGCTCTGCC ACTCAGGGCG | 1680 |
| TGGTAAGATC CCACTTCTTG TTCACTTCGG ATAATGGCCT AATTACTATT GCAGGGTGGTA | 1740 |
| AATGGACTAC TTACAGACAA ATGGCTGAGG AAACAGTCGA CAAAGTTGTC GAAGTTGGCG | 1800 |
| GATTCCACAA CCTGAAACCT TGTCACACAA GAGATATTAA GCTTGCTGGT GCAGAAGAAT | 1860 |
| GGACGCAAAA CTATGTGGCT TTATTGGCTC AAAACTACCA TTTATCATCA AAAATGTCCA | 1920 |
| ACTACTTGGT TCAAAACTAC GGAACCCGTT CCTCTATCAT TTGCGAATTT TTCAAAGAAT | 1980 |
| CCATGGAAAA TAAACTGCCT TTGTCCTTAG CCGACAAGGA AAATAACGTA ATCTACTCTA | 2040 |
| GCGAGGAGAA CAACTTGGTC AATTTGATA CTTTCAGATA TCCATTACAA ATCGGTGAGT | 2100 |
| TAAAGTATTTC CATGCAGTAC GAATATTGTA GAACTCCCTT GGACTTCCTT TTAAGAAGAA | 2160 |
| CAAGATTCGC CTTCTTGGAC GCCAAGGAAG CTTTGAATGC CGTGCATGCC ACCGTCAAAG | 2220 |
| TTATGGGTGA TGAGTTCAAT TGGTCGGAGA AAAAGAGGCA GTGGGAACCTT GAAAAAAACTG | 2280 |
| TGAACTTCAT CCAAGGACGT TTCGGTGTCT AAATCGATCA TGATAGTTAA GGGTGACAAA | 2340 |
| GATAACATTC ACAAGAGTAA TAATAATGGT AATGATGATA ATAATAATAA TGATAGTAAT | 2400 |
| AACAATAATA ATAATGGTGG TAATGGCAAT GAAATCGCTA TTATTACCTA TTTTCCTTAA | 2460 |
| TGGAAGAGTT AAAGTAAACT AAAAAAACTA CAAAAATATA TGAAGAAAAA AAAAAAAAGA | 2520 |
| GGTAATAGAC TCTACTACTA CAATTGATCT TCAAATTATG ACCTTCCTAG TGTTTATATT | 2580 |
| CTATTTCAA TACATAATAT AATCTATATA ATCATTGCTG GTAGACTTCC GTTTAATAT | 2640 |
| CGTTTTAATT ATCCCCTTA TCTCTAGTCT AGTTTTATCA TAAAATATAG AAACACTAAA | 2700 |
| TAATATTCTT CAAACGGTCC TGGTGCATAC GCAATACATA TTTATGGTGC AAAAAAAA | 2760 |
| ATGGAAAATT TTGCTAGTCA TAAACCCTTT CATAAAACAA TACGTAGACA TCGCTACTTG | 2820 |
| AAATTTCAA GTTTTATCA GATCCATGTT TCCTATCTGC CTTGACAACC TCATCGTCGA | 2880 |
| AATAGTACCA TTTAGAACGC CCAATATTCA CATTGTGTTA AAGGTCTTAA TTCACCAGTG | 2940 |
| ACGTGTAATG GCCATGATTA ATGTGCCTGT ATGGTTAACCTC ACTCCAAATA GCTTATATT | 3000 |
| CATAGTGTCA TTGTTTTCA ATATAATGTT TAGTATCAAT GGATATGTTA CGACGGTGT | 3060 |
| ATTTTTCTTG GTCAAATCGT AATAAAATCT CGATAAAATGG ATGACTAAGA TTTTTGGTAA | 3120 |
| AGTTACAAAAA TTTATCGTT TCACTGTTGT CAATTTTTG TTCTTGTAAAT CACTCGAG | 3178 |

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 816 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

| | |
|---|-----|
| ATGAAACGTT TCAATGTTT AAAATATATC AGAACACAAA AAGCAAATAT ACAAAACCATC | 60 |
| GCAATGCCTT TGACCACAAA ACCTTATCT TTGAAAATCA ACGCCGCTCT ATTGATGTT | 120 |
| GACGGTACCA TCATCATCTC TCAACCAGCC ATTGCTGCTT TCTGGAGAGA TTTCGGTAAA | 180 |
| GACAAGCCTT ACTTCGATGC CGAACACGTT ATTACATCT CTCACGGTTG GAGAACATTAC | 240 |
| GATGCCATTG CCAAGTCGC TCCAGACTTT GCTGATGAAG AATACGTTAA CAAGCTAGAA | 300 |
| GGTGAAATCC CAGAAAAGTA CGGTGAACAC TCCATCGAAG TTCCAGGTGC TGTCAAGTTG | 360 |
| TGTAATGCTT TGAACGCCTT GCCAAAGGAA AAATGGGCTG TCGCCACCTC TGGTACCCGT | 420 |
| GACATGGCCA AGAAATGGTT CGACATTTG AAGATCAAGA GACCAGAATA CTTCATCACC | 480 |
| GCCAATGATG TCAAGCAAGG TAAGCCTCAC CCAGAACCAT ACTTAAAGGG TAGAACGGT | 540 |
| TTGGGTTTCC CAATTAATGA ACAAGACCCA TCCAAATCTA AGGTTGTTGT CTTGAAGAC | 600 |
| GCACCAGCTG GTATTGCTGC TGGTAAGGCT GCTGGCTGTA AAATCGTTGG TATTGCTACC | 660 |
| ACTTCGATT TGGACTTCTT GAAGGAAAAG GGTTGTGACA TCATTGTCAA GAACCACGAA | 720 |
| TCTATCAGAG TCGGTGAATA CAACGCTGAA ACCGATGAAG TCGAATTGAT CTTTGATGAC | 780 |
| TAATTATACG CTAAGGATGA CTTGTTGAAA TGGTAA | 816 |

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 753 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

| | |
|---|-----|
| ATGGGATTGA CTACTAAACC TCTATCTTG AAAGTTAACG CCGCTTGTT CGACGTCGAC | 60 |
| GGTACCATTA TCATCTCTCA ACCAGCCATT GCTGCATTCT GGAGGGATTT CGGTAAGGAC | 120 |
| AAACCTTATT TCGATGCTGA ACACGTTATC CAAGTCTCGC ATGGTTGGAG AACGTTTGAT | 180 |

GCCATTGCTA AGTCGCTCC AGACTTGCC AATGAAGAGT ATGTTAACAA ATTAGAAC 240
 GAAATTCCGG TCAAGTACGG TGAAAAATCC ATTGAAGTCC CAGGTGCAGT TAAGCTGTGC 300
 AACGCTTGTA ACGCTCTACC AAAAGAGAAA TGGGCTGTGG CAACTTCCGG TACCCGTGAT 360
 ATGGCACAAA AATGGTTCGA GCATCTGGGA ATCAGGAGAC CAAAGTACTT CATTACCGCT 420
 AATGATGTCA AACAGGGTAA GCCTCATCCA GAACCATATC TGAAGGGCAG GAATGGCTTA 480
 GGATATCCGA TCAATGAGCA AGACCCTTCC AAATCTAAGG TAGTAGTATT TGAAGACGCT 540
 CCAGCAGGTA TTGCCGCCGG AAAAGCCGCC GGTTGTAAGA TCATTGGTAT TGCCACTACT 600
 TTCGACTTGG ACTTCCTAAA GGAAAAAGGC TGTGACATCA TTGTCAAAAA CCACGAATCC 660
 ATCAGAGTTG GCGGCTACAA TGCCGAAACA GACGAAGTTG AATTCACTTT TGACGACTAC 720
 TTATATGCTA AGGACGATCT GTTGAAATGG TAA 753

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2520 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TGTATTGGCC ACGATAACCA CCCTTGTAT ACTGTTTTG TTTTCACAT GGTAAATAAC 60
 GACTTTTATT AAACAACGTA TGAAAAACA TAACAAGAAT CTACCCATAC AGGCCATTTC 120
 GTAATTCTTC TCTTCTAATT GGAGTAAAAC CATCAATTAA AGGGTGTGGA GTAGCATAGT 180
 GAGGGGCTGA CTGCATTGAC AAAAAAATTG AAAAAAAAAA AGGAAAAGGA AAGGAAAAAA 240
 AGACAGCCAA GACTTTAGA ACGGATAAGG TGTAATAAAA TGTGGGGGGA TGCCTGTTCT 300
 CGAACCATAT AAAATATACC ATGTGGTTG AGTTGTGCC GGAACATAC AAATAGTTAT 360
 ATGTTCCCT CTCTCTTCCG ACTTGTAGTA TTCTCCAAAC GTTACATATT CCGATCAAGC 420
 CAGCGCCTT ACACTAGTTT AAAACAAGAA CAGAGCCGTA TGTCCAAAAT AATGGAAGAT 480
 TTACGAAGTG ACTACGTCCC GCTTATGCC AGTATTGATG TAGGAACGAC CTCATCCAGA 540
 TGCATTCTGT TCAACAGATG GGGCCAGGAC GTTCAAAAC ACCAAATTGA ATATTCAACT 600
 TCAGCATCGA AGGGCAAGAT TGGGGTGTCT GGCCTAAGGA GACCCCTCTAC AGCCCCAGCT 660
 CGTGAACAC CAAACGCCGG TGACATCAA ACCAGCGGAA AGCCCATCTT TTCTGCAGAA 720
 GGCTATGCCA TTCAAGAAC CAAATTCTA AAAATCGAGG AATTGGACTT GGACTTCCAT 780

AACGAACCCA CGTTGAAGTT CCCCAAACCG GGTTGGGTTG AGTGCATCC GCAGAAATT 840
 CTGGTGAACG TCGTCCAATG CCTTGCTCA AGTTTGCTCT CTCTGCAGAC TATCAACAGC 900
 GAACGTGTAG CAAACGGTCT CCCACCTTAC AAGGTAATAT GCATGGGTAT AGCAAACATG 960
 AGAGAAACCA CAATTCTGTG GTCCCCGCCGC ACAGGAAAAC CAATTGTTAA CTACGGTATT 1020
 GTTTGGAACG ACACCAGAAC GATCAAAATC GTTAGAGACA AATGGCAAAA CACTAGCGTC 1080
 GATAGGCAAC TGCAGCTTAG ACAGAAGACT GGATTGCCAT TGCTCTCCAC GTATTTCTCC 1140
 TGTTCCAAGC TGGCCTGGTT CCTCGACAAT GAGCCTCTGT GTACCAAGGC GTATGAGGAG 1200
 AACGACCTGA TGTCGGCAC TGTGGACACA TGGCTGATT ACCAATTAAC TAAACAAAAG 1260
 GCGTCGTTT CTGACGTAAC CAACGCTTCC AGAACTGGAT TTATGAACCT CTCCACTTTA 1320
 AAGTACGACA ACGAGTTGCT GGAATTTGG GGTATTGACA AGAACCTGAT TCACATGCC 1380
 GAAATTGTGT CCTCATCTCA ATACTACGGT GACTTTGGCA TTCCTGATTG GATAATGGAA 1440
 AAGCTACACG ATTGCCCCAA AACAGTACTG CGAGATCTAG TCAAGAGAAA CCTGCCATA 1500
 CAGGGCTGTC TGGGCGACCA AAGCGCATCC ATGGTGGGC AACTCGCTTA CAAACCCGGT 1560
 GCTGCAAAAT GTACTTATGG TACCGGTTGC TTTTACTGT ACAATACGGG GACCAAAAAA 1620
 TTGATCTCCC AACATGGCGC ACTGACGACT CTAGCATTT GGTTCCCACA TTTGCAAGAG 1680
 TACGGTGGCC AAAAACCAGA ATTGAGCAAG CCACATTTG CATTAGAGGG TTCCGTCGCT 1740
 GTGGCTGGTG CTGTGGTCCA ATGGCTACGT GATAATTAC GATTGATCGA TAAATCAGAG 1800
 GATGTCGGAC CGATTGCATC TACGGTTCCCT GATTCTGGTG GCGTAGTTT CGTCCCCGCA 1860
 TTTAGTGGCC TATTGCTCC CTATTGGAC CCAGATGCCA GAGCCACCAT AATGGGGATG 1920
 TCTCAATTCA CTACTGCCTC CCACATGCC AGAGCTGCCG TGGAAGGTGT TTGCTTCAA 1980
 GCCAGGGCTA TCTTGAGGC AATGAGTTCT GACCGCTTG GTGAAGGTTC CAAAGACAGG 2040
 GACTTTTAG AGGAAATTTC CGACGTCACA TATGAAAAGT CGCCCCGTGTC GGTTCTGGCA 2100
 GTGGATGGCG GGATGTCGAG GTCTAATGAA GTCATGCAA TTCAAGCCGA TATCCTAGGT 2160
 CCCTGTGTCA AAGTCAGAAC GTCTCCGACA GCGGAATGTA CCGCATTGGG GGCAGCCATT 2220
 GCAGCCAATA TGGCTTCAA GGATGTGAAC GAGCGCCCAT TATGGAAGGA CCTACACGAT 2280
 GTTAAGAAAT GGGCTTTTA CAATGGAATG GAGAAAAACG AACAAATATC ACCAGAGGCT 2340
 CATCCAAACC TTAAGATATT CAGAAGTGAA TCCGACGATG CTGAAAGGAG AAAGCATTGG 2400
 AAGTATTGGG AAGTTGCCGT GGAAAGATCC AAAGGTTGGC TGAAGGACAT AGAAGGTGAA 2460
 CACGAACAGG TTCTAGAAAA CTTCCAATAA CAACATAAAT AATTCTATT AACAAATGTAA 2520

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 391 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ser | Ala | Ala | Ala | Asp | Arg | Leu | Asn | Leu | Thr | Ser | Gly | His | Leu | Asn |
| 1 | | | | | | | 5 | | | 10 | | | | 15 | |
| | | | | | | | | | | | | | | | |
| Ala | Gly | Arg | Lys | Arg | Ser | Ser | Ser | Val | Ser | Leu | Lys | Ala | Ala | Glu | |
| | | | | | | | | 20 | | 25 | | | | 30 | |
| | | | | | | | | | | | | | | | |
| Lys | Pro | Phe | Lys | Val | Thr | Val | Ile | Gly | Ser | Gly | Asn | Trp | Gly | Thr | Thr |
| | | | | | | | 35 | | 40 | | | | 45 | | |
| | | | | | | | | | | | | | | | |
| Ile | Ala | Lys | Val | Val | Ala | Glu | Asn | Cys | Lys | Gly | Tyr | Pro | Glu | Val | Phe |
| | | | | | | 50 | | 55 | | | 60 | | | | |
| | | | | | | | | | | | | | | | |
| Ala | Pro | Ile | Val | Gln | Met | Trp | Val | Phe | Glu | Glu | Glu | Ile | Asn | Gly | Glu |
| | | | | | | | 65 | 70 | | 75 | | | | 80 | |
| | | | | | | | | | | | | | | | |
| Lys | Leu | Thr | Glu | Ile | Ile | Asn | Thr | Arg | His | Gln | Asn | Val | Lys | Tyr | Leu |
| | | | | | | | 85 | | 90 | | | | 95 | | |
| | | | | | | | | | | | | | | | |
| Pro | Gly | Ile | Thr | Leu | Pro | Asp | Asn | Leu | Val | Ala | Asn | Pro | Asp | Leu | Ile |
| | | | | | | | 100 | | 105 | | | | 110 | | |
| | | | | | | | | | | | | | | | |
| Asp | Ser | Val | Lys | Asp | Val | Asp | Ile | Ile | Val | Phe | Asn | Ile | Pro | His | Gln |
| | | | | | | | 115 | | 120 | | | | 125 | | |
| | | | | | | | | | | | | | | | |
| Phe | Leu | Pro | Arg | Ile | Cys | Ser | Gln | Leu | Lys | Gly | His | Val | Asp | Ser | His |
| | | | | | | | 130 | | 135 | | | | 140 | | |
| | | | | | | | | | | | | | | | |
| Val | Arg | Ala | Ile | Ser | Cys | Leu | Lys | Gly | Phe | Glu | Val | Gly | Ala | Lys | Gly |
| | | | | | | | 145 | | 150 | | 155 | | | 160 | |
| | | | | | | | | | | | | | | | |
| Val | Gln | Leu | Leu | Ser | Ser | Tyr | Ile | Thr | Glu | Glu | Leu | Gly | Ile | Gln | Cys |
| | | | | | | | 165 | | 170 | | | | 175 | | |
| | | | | | | | | | | | | | | | |
| Gly | Ala | Leu | Ser | Gly | Ala | Asn | Ile | Ala | Thr | Glu | Val | Ala | Gln | Glu | His |
| | | | | | | | 180 | | 185 | | | | 190 | | |
| | | | | | | | | | | | | | | | |
| Trp | Ser | Glu | Thr | Thr | Val | Ala | Tyr | His | Ile | Pro | Lys | Asp | Phe | Arg | Gly |
| | | | | | | | 195 | | 200 | | | | 205 | | |
| | | | | | | | | | | | | | | | |
| Glu | Gly | Lys | Asp | Val | Asp | His | Lys | Val | Leu | Lys | Ala | Leu | Phe | His | Arg |
| | | | | | | | 210 | | 215 | | | | 220 | | |
| | | | | | | | | | | | | | | | |
| Pro | Tyr | Phe | His | Val | Ser | Val | Ile | Glu | Asp | Val | Ala | Gly | Ile | Ser | Ile |
| | | | | | | | 225 | | 230 | | 235 | | | 240 | |

Cys Gly Ala Leu Lys Asn Val Val Ala Leu Gly Cys Gly Phe Val Glu
 245 250 255
 Gly Leu Gly Trp Gly Asn Asn Ala Ser Ala Ala Ile Gln Arg Val Gly
 260 265 270
 Leu Gly Glu Ile Ile Arg Phe Gly Gln Met Phe Phe Pro Glu Ser Arg
 275 280 285
 Glu Glu Thr Tyr Tyr Gln Glu Ser Ala Gly Val Ala Asp Leu Ile Thr
 290 295 300
 Thr Cys Ala Gly Gly Arg Asn Val Lys Val Ala Arg Leu Met Ala Thr
 305 310 315 320
 Ser Gly Lys Asp Ala Trp Glu Cys Glu Lys Glu Leu Leu Asn Gly Gln
 325 330 335
 Ser Ala Gln Gly Leu Ile Thr Cys Lys Glu Val His Glu Trp Leu Glu
 340 345 350
 Thr Cys Gly Ser Val Glu Asp Phe Pro Leu Phe Glu Ala Val Tyr Gln
 355 360 365
 Ile Val Tyr Asn Asn Tyr Pro Met Lys Asn Leu Pro Asp Met Ile Glu
 370 375 380
 Glu Leu Asp Leu His Glu Asp
 385 390

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 384 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Thr Ala His Thr Asn Ile Lys Gln His Lys His Cys His Glu Asp
 1 5 10 15
 His Pro Ile Arg Arg Ser Asp Ser Ala Val Ser Ile Val His Leu Lys
 20 25 30
 Arg Ala Pro Phe Lys Val Thr Val Ile Gly Ser Gly Asn Trp Gly Thr
 35 40 45
 Thr Ile Ala Lys Val Ile Ala Glu Asn Thr Glu Leu His Ser His Ile
 50 55 60
 Phe Glu Pro Glu Val Arg Met Trp Val Phe Asp Glu Lys Ile Gly Asp
 65 70 75 80

Glu Asn Leu Thr Asp Ile Ile Asn Thr Arg His Gln Asn Val Lys Tyr
 85 90 95
 Leu Pro Asn Ile Asp Leu Pro His Asn Leu Val Ala Asp Pro Asp Leu
 100 105 110
 Leu His Ser Ile Lys Gly Ala Asp Ile Leu Val Phe Asn Ile Pro His
 115 120 125
 Gln Phe Leu Pro Asn Ile Val Lys Gln Leu Gln Gly His Val Ala Pro
 130 135 140
 His Val Arg Ala Ile Ser Cys Leu Lys Gly Phe Glu Leu Gly Ser Lys
 145 150 155 160
 Gly Val Gln Leu Leu Ser Ser Tyr Val Thr Asp Glu Leu Gly Ile Gln
 165 170 175
 Cys Gly Ala Leu Ser Gly Ala Asn Leu Ala Pro Glu Val Ala Lys Glu
 180 185 190
 His Trp Ser Glu Thr Thr Val Ala Tyr Gln Leu Pro Lys Asp Tyr Gln
 195 200 205
 Gly Asp Gly Lys Asp Val Asp His Lys Ile Leu Lys Leu Leu Phe His
 210 215 220
 Arg Pro Tyr Phe His Val Asn Val Ile Asp Asp Val Ala Gly Ile Ser
 225 230 235 240
 Ile Ala Gly Ala Leu Lys Asn Val Val Ala Leu Ala Cys Gly Phe Val
 245 250 255
 Glu Gly Met Gly Trp Gly Asn Asn Ala Ser Ala Ile Gln Arg Leu
 260 265 270
 Gly Leu Gly Glu Ile Ile Lys Phe Gly Arg Met Phe Phe Pro Glu Ser
 275 280 285
 Lys Val Glu Thr Tyr Tyr Gln Glu Ser Ala Gly Val Ala Asp Leu Ile
 290 295 300
 Thr Thr Cys Ser Gly Gly Arg Asn Val Lys Val Ala Thr Tyr Met Ala
 305 310 315 320
 Lys Thr Gly Lys Ser Ala Leu Glu Ala Glu Lys Glu Leu Leu Asn Gly
 325 330 335
 Gln Ser Ala Gln Gly Ile Ile Thr Cys Arg Glu Val His Glu Trp Leu
 340 345 350
 Gln Thr Cys Glu Leu Thr Gln Glu Phe Pro Ile Ile Arg Gly Ser Leu
 355 360 365
 Pro Asp Ser Leu Gln Gln Arg Pro His Gly Arg Pro Thr Gly Asp Asp
 370 375 380

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 614 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Thr | Arg | Ala | Thr | Trp | Cys | Asn | Ser | Pro | Pro | Pro | Leu | His | Arg | Gln |
| 1 | | | | | 5 | | | | 10 | | | | 15 | | |
| Val | Ser | Arg | Arg | Asp | Leu | Leu | Asp | Arg | Leu | Asp | Lys | Thr | His | Gln | Phe |
| | | | | 20 | | | | 25 | | | | 30 | | | |
| Asp | Val | Leu | Ile | Ile | Gly | Gly | Ala | Thr | Gly | Thr | Gly | Cys | Ala | Leu | |
| | | | | 35 | | | | 40 | | | | 45 | | | |
| Asp | Ala | Ala | Thr | Arg | Gly | Leu | Asn | Val | Ala | Leu | Val | Glu | Lys | Gly | Asp |
| | | | | 50 | | | | 55 | | | | 60 | | | |
| Phe | Ala | Ser | Gly | Thr | Ser | Ser | Lys | Ser | Thr | Lys | Met | Ile | His | Gly | Gly |
| | | | | 65 | | | | 70 | | | 75 | | | 80 | |
| Val | Arg | Tyr | Leu | Glu | Lys | Ala | Phe | Trp | Glu | Phe | Ser | Lys | Ala | Gln | Leu |
| | | | | 85 | | | | 90 | | | | | 95 | | |
| Asp | Leu | Val | Ile | Glu | Ala | Leu | Asn | Glu | Arg | Lys | His | Leu | Ile | Asn | Thr |
| | | | | 100 | | | | 105 | | | | 110 | | | |
| Ala | Pro | His | Leu | Cys | Thr | Val | Leu | Pro | Ile | Leu | Ile | Pro | Ile | Tyr | Ser |
| | | | | 115 | | | | 120 | | | | 125 | | | |
| Thr | Trp | Gln | Val | Pro | Tyr | Ile | Tyr | Met | Gly | Cys | Lys | Phe | Tyr | Asp | Phe |
| | | | | 130 | | | | 135 | | | | 140 | | | |
| Phe | Gly | Gly | Ser | Gln | Asn | Leu | Lys | Lys | Ser | Tyr | Leu | Leu | Ser | Lys | Ser |
| | | | | 145 | | | | 150 | | | 155 | | | 160 | |
| Ala | Thr | Val | Glu | Lys | Ala | Pro | Met | Leu | Thr | Thr | Asp | Asn | Leu | Lys | Ala |
| | | | | 165 | | | | 170 | | | | 175 | | | |
| Ser | Leu | Val | Tyr | His | Asp | Gly | Ser | Phe | Asn | Asp | Ser | Arg | Leu | Asn | Ala |
| | | | | 180 | | | | 185 | | | | 190 | | | |
| Thr | Leu | Ala | Ile | Thr | Gly | Val | Glu | Asn | Gly | Ala | Thr | Val | Leu | Ile | Tyr |
| | | | | 195 | | | | 200 | | | | 205 | | | |
| Val | Glu | Val | Gln | Lys | Leu | Ile | Lys | Asp | Pro | Thr | Ser | Gly | Lys | Val | Ile |
| | | | | 210 | | | | 215 | | | | 220 | | | |
| Gly | Ala | Glu | Ala | Arg | Asp | Val | Glu | Thr | Asn | Glu | Leu | Val | Arg | Ile | Asn |
| | | | | 225 | | | | 230 | | | 235 | | | 240 | |

Ala Lys Cys Val Val Asn Ala Thr Gly Pro Tyr Ser Asp Ala Ile Leu
 245 250 255
 Gln Met Asp Arg Asn Pro Ser Gly Leu Pro Asp Ser Pro Leu Asn Asp
 260 265 270
 Asn Ser Lys Ile Lys Ser Thr Phe Asn Gln Ile Ser Val Met Asp Pro
 275 280 285
 Lys Met Val Ile Pro Ser Ile Gly Val His Ile Val Leu Pro Ser Phe
 290 295 300
 Tyr Ser Pro Lys Asp Met Gly Leu Leu Asp Val Arg Thr Ser Asp Gly
 305 310 315 320
 Arg Val Met Phe Phe Leu Pro Trp Gln Gly Lys Val Leu Ala Gly Thr
 325 330 335
 Thr Asp Ile Pro Leu Lys Gln Val Pro Glu Asn Pro Met Pro Thr Glu
 340 345 350
 Ala Asp Ile Gln Asp Ile Leu Lys Glu Leu Gln His Tyr Ile Glu Phe
 355 360 365
 Pro Val Lys Arg Glu Asp Val Leu Ser Ala Trp Ala Gly Val Arg Pro
 370 375 380
 Leu Val Arg Asp Pro Arg Thr Ile Pro Ala Asp Gly Lys Lys Gly Ser
 385 390 395 400
 Ala Thr Gln Gly Val Val Arg Ser His Phe Leu Phe Thr Ser Asp Asn
 405 410 415
 Gly Leu Ile Thr Ile Ala Gly Gly Lys Trp Thr Thr Tyr Arg Gln Met
 420 425 430
 Ala Glu Glu Thr Val Asp Lys Val Val Glu Val Gly Gly Phe His Asn
 435 440 445
 Leu Lys Pro Cys His Thr Arg Asp Ile Lys Leu Ala Gly Ala Glu Glu
 450 455 460
 Trp Thr Gln Asn Tyr Val Ala Leu Leu Ala Gln Asn Tyr His Leu Ser
 465 470 475 480
 Ser Lys Met Ser Asn Tyr Leu Val Gln Asn Tyr Gly Thr Arg Ser Ser
 485 490 495
 Ile Ile Cys Glu Phe Phe Lys Glu Ser Met Glu Asn Lys Leu Pro Leu
 500 505 510
 Ser Leu Ala Asp Lys Glu Asn Asn Val Ile Tyr Ser Ser Glu Glu Asn
 515 520 525
 Asn Leu Val Asn Phe Asp Thr Phe Arg Tyr Pro Phe Thr Ile Gly Glu
 530 535 540

Leu Lys Tyr Ser Met Gln Tyr Glu Tyr Cys Arg Thr Pro Leu Asp Phe
 545 550 555 560
 Leu Leu Arg Arg Thr Arg Phe Ala Phe Leu Asp Ala Lys Glu Ala Leu
 565 570 575
 Asn Ala Val His Ala Thr Val Lys Val Met Gly Asp Glu Phe Asn Trp
 580 585 590
 Ser Glu Lys Lys Arg Gln Trp Glu Leu Glu Lys Thr Val Asn Phe Ile
 595 600 605
 Gln Gly Arg Phe Gly Val
 610

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 339 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

- (ii) MOLECULE TYPE: protein

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Asn Gln Arg Asn Ala Ser Met Thr Val Ile Gly Ala Gly Ser Tyr
 1 5 10 15
 Gly Thr Ala Leu Ala Ile Thr Leu Ala Arg Asn Gly His Glu Val Val
 20 25 30
 Leu Trp Gly His Asp Pro Glu His Ile Ala Thr Leu Glu Arg Asp Arg
 35 40 45
 Cys Asn Ala Ala Phe Leu Pro Asp Val Pro Phe Pro Asp Thr Leu His
 50 55 60
 Leu Glu Ser Asp Leu Ala Thr Ala Leu Ala Ala Ser Arg Asn Ile Leu
 65 70 75 80
 Val Val Val Pro Ser His Val Phe Gly Glu Val Leu Arg Gln Ile Lys
 85 90 95
 Pro Leu Met Arg Pro Asp Ala Arg Leu Val Trp Ala Thr Lys Gly Leu
 100 105 110
 Glu Ala Glu Thr Gly Arg Leu Leu Gln Asp Val Ala Arg Glu Ala Leu
 115 120 125
 Gly Asp Gln Ile Pro Leu Ala Val Ile Ser Gly Pro Thr Phe Ala Lys
 130 135 140
 Glu Leu Ala Ala Gly Leu Pro Thr Ala Ile Ser Leu Ala Ser Thr Asp
 145 150 155 160

Gln Thr Phe Ala Asp Asp Leu Gln Gln Leu Leu His Cys Gly Lys Ser
165 170 175

Phe Arg Val Tyr Ser Asn Pro Asp Phe Ile Gly Val Gln Leu Gly Gly
180 185 190

Ala Val Lys Asn Val Ile Ala Ile Gly Ala Gly Met Ser Asp Gly Ile
195 200 205

Gly Phe Gly Ala Asn Ala Arg Thr Ala Leu Ile Thr Arg Gly Leu Ala
210 215 220

Glu Met Ser Arg Leu Gly Ala Ala Leu Gly Ala Asp Pro Ala Thr Phe
225 230 235 240

Met Gly Met Ala Gly Leu Gly Asp Leu Val Leu Thr Cys Thr Asp Asn
245 250 255

Gln Ser Arg Asn Arg Arg Phe Gly Met Met Leu Gly Gln Gly Met Asp
260 265 270

Val Gln Ser Ala Gln Glu Lys Ile Gly Gln Val Val Glu Gly Tyr Arg
275 280 285

Asn Thr Lys Glu Val Arg Glu Leu Ala His Arg Phe Gly Val Glu Met
290 295 300

Pro Ile Thr Glu Glu Ile Tyr Gln Val Leu Tyr Cys Gly Lys Asn Ala
305 310 315 320

Arg Glu Ala Ala Leu Thr Leu Leu Gly Arg Ala Arg Lys Asp Glu Arg
325 330 335

Ser Ser His

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 501 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Glu Thr Lys Asp Leu Ile Val Ile Gly Gly Ile Asn Gly Ala
1 5 10 15

Gly Ile Ala Ala Asp Ala Ala Gly Arg Gly Leu Ser Val Leu Met Leu
20 25 30

Glu Ala Gln Asp Leu Ala Cys Ala Thr Ser Ser Ala Ser Ser Lys Leu
35 40 45

Ile His Gly Gly Leu Arg Tyr Leu Glu His Tyr Glu Phe Arg Leu Val
50 55 60

Ser Glu Ala Leu Ala Glu Arg Glu Val Leu Leu Lys Met Ala Pro His
 65 .70 75 80
 Ile Ala Phe Pro Met Arg Phe Arg Leu Pro His Arg Pro His Leu Arg
 85 90 95
 Pro Ala Trp Met Ile Arg Ile Gly Leu Phe Met Tyr Asp His Leu Gly
 100 105 110
 Lys Arg Thr Ser Leu Pro Gly Ser Thr Gly Leu Arg Phe Gly Ala Asn
 115 120 125
 Ser Val Leu Lys Pro Glu Ile Lys Arg Gly Phe Glu Tyr Ser Asp Cys
 130 135 140
 Trp Val Asp Asp Ala Arg Leu Val Leu Ala Asn Ala Gln Met Val Val
 145 150 155 160
 Arg Lys Gly Gly Glu Val Leu Thr Arg Thr Arg Ala Thr Ser Ala Arg
 165 170 175
 Arg Glu Asn Gly Leu Trp Ile Val Glu Ala Glu Asp Ile Asp Thr Gly
 180 185 190
 Lys Lys Tyr Ser Trp Gln Ala Arg Gly Leu Val Asn Ala Thr Gly Pro
 195 200 205
 Trp Val Lys Gln Phe Phe Asp Asp Gly Met His Leu Pro Ser Pro Tyr
 210 215 220
 Gly Ile Arg Leu Ile Lys Gly Ser His Ile Val Val Pro Arg Val His
 225 230 235 240
 Thr Gln Lys Gln Ala Tyr Ile Leu Gln Asn Glu Asp Lys Arg Ile Val
 245 250 255
 Phe Val Ile Pro Trp Met Asp Glu Phe Ser Ile Ile Gly Thr Thr Asp
 260 265 270
 Val Glu Tyr Lys Gly Asp Pro Lys Ala Val Lys Ile Glu Glu Ser Glu
 275 280 285
 Ile Asn Tyr Leu Leu Asn Val Tyr Asn Thr His Phe Lys Lys Gln Leu
 290 295 300
 Ser Arg Asp Asp Ile Val Trp Thr Tyr Ser Gly Val Arg Pro Leu Cys
 305 310 315 320
 Asp Asp Glu Ser Asp Ser Pro Gln Ala Ile Thr Arg Asp Tyr Thr Leu
 325 330 335
 Asp Ile His Asp Glu Asn Gly Lys Ala Pro Leu Leu Ser Val Phe Gly
 340 345 350
 Gly Lys Leu Thr Thr Tyr Arg Lys Leu Ala Glu His Ala Leu Glu Lys
 355 360 365

Leu Thr Pro Tyr Tyr Gln Gly Ile Gly Pro Ala Trp Thr Lys Glu Ser
 370 375 380
 Val Leu Pro Gly Gly Ala Ile Glu Gly Asp Arg Asp Asp Tyr Ala Ala
 385 390 395 400
 Arg Leu Arg Arg Arg Tyr Pro Phe Leu Thr Glu Ser Leu Ala Arg His
 405 410 415
 Tyr Ala Arg Thr Tyr Gly Ser Asn Ser Glu Leu Leu Gly Asn Ala
 420 425 430
 Gly Thr Val Ser Asp Leu Gly Glu Asp Phe Gly His Glu Phe Tyr Glu
 435 440 445
 Ala Glu Leu Lys Tyr Leu Val Asp His Glu Trp Val Arg Arg Ala Asp
 450 455 460
 Asp Ala Leu Trp Arg Arg Thr Lys Gln Gly Met Trp Leu Asn Ala Asp
 465 470 475 480
 Gln Gln Ser Arg Val Ser Gln Trp Leu Val Glu Tyr Thr Gln Gln Arg
 485 490 495
 Leu Ser Leu Ala Ser
 500

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 542 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Lys Thr Arg Asp Ser Gln Ser Ser Asp Val Ile Ile Ile Gly Gly
 1 5 10 15
 Gly Ala Thr Gly Ala Gly Ile Ala Arg Asp Cys Ala Leu Arg Gly Leu
 20 25 30
 Arg Val Ile Leu Val Glu Arg His Asp Ile Ala Thr Gly Ala Thr Gly
 35 40 45
 Arg Asn His Gly Leu Leu His Ser Gly Ala Arg Tyr Ala Val Thr Asp
 50 55 60
 Ala Glu Ser Ala Arg Glu Cys Ile Ser Glu Asn Gln Ile Leu Lys Arg
 65 70 75 80
 Ile Ala Arg His Cys Val Glu Pro Thr Asn Gly Leu Phe Ile Thr Leu
 85 90 95

Pro Glu Asp Asp Leu Ser Phe Gln Ala Thr Phe Ile Arg Ala Cys Glu
100 105 110

Glu Ala Gly Ile Ser Ala Glu Ala Ile Asp Pro Gln Gln Ala Arg Ile
115 120 125

Ile Glu Pro Ala Val Asn Pro Ala Leu Ile Gly Ala Val Lys Val Pro
130 135 140

Asp Gly Thr Val Asp Pro Phe Arg Leu Thr Ala Ala Asn Met Leu Asp
145 150 155 160

Ala Lys Glu His Gly Ala Val Ile Leu Thr Ala His Glu Val Thr Gly
165 170 175

Leu Ile Arg Glu Gly Ala Thr Val Cys Gly Val Arg Val Arg Asn His
180 185 190

Leu Thr Gly Glu Thr Gln Ala Leu His Ala Pro Val Val Asn Ala
195 200 205

Ala Gly Ile Trp Gly Gln His Ile Ala Glu Tyr Ala Asp Leu Arg Ile
210 215 220

Arg Met Phe Pro Ala Lys Gly Ser Leu Leu Ile Met Asp His Arg Ile
225 230 235 240

Asn Gln His Val Ile Asn Arg Cys Arg Lys Pro Ser Asp Ala Asp Ile
245 250 255

Leu Val Pro Gly Asp Thr Ile Ser Leu Ile Gly Thr Thr Ser Leu Arg
260 265 270

Ile Asp Tyr Asn Glu Ile Asp Asp Asn Arg Val Thr Ala Glu Glu Val
275 280 285

Asp Ile Leu Leu Arg Glu Gly Glu Lys Leu Ala Pro Val Met Ala Lys
290 295 300

Thr Arg Ile Leu Arg Ala Tyr Ser Gly Val Arg Pro Leu Val Ala Ser
305 310 315 320

Asp Asp Asp Pro Ser Gly Arg Asn Leu Ser Arg Gly Ile Val Leu Leu
325 330 335

Asp His Ala Glu Arg Asp Gly Leu Asp Gly Phe Ile Thr Ile Thr Gly
340 345 350

Gly Lys Leu Met Thr Tyr Arg Leu Met Ala Glu Trp Ala Thr Asp Ala
355 360 365

Val Cys Arg Lys Leu Gly Asn Thr Arg Pro Cys Thr Thr Ala Asp Leu
370 375 380

Ala Leu Pro Gly Ser Gln Glu Pro Ala Glu Val Thr Leu Arg Lys Val
385 390 395 400

Ile Ser Leu Pro Ala Pro Leu Arg Gly Ser Ala Val Tyr Arg His Gly
 405 410 415
 Asp Arg Thr Pro Ala Trp Leu Ser Glu Gly Arg Leu His Arg Ser Leu
 420 425 430
 Val Cys Glu Cys Glu Ala Val Thr Ala Gly Glu Val Gln Tyr Ala Val
 435 440 445
 Glu Asn Leu Asn Val Asn Ser Leu Leu Asp Leu Arg Arg Arg Thr Arg
 450 455 460
 Val Gly Met Gly Thr Cys Gln Gly Glu Leu Cys Ala Cys Arg Ala Ala
 465 470 475 480
 Gly Leu Leu Gln Arg Phe Asn Val Thr Thr Ser Ala Gln Ser Ile Glu
 485 490 495
 Gln Leu Ser Thr Phe Leu Asn Glu Arg Trp Lys Gly Val Gln Pro Ile
 500 505 510
 Ala Trp Gly Asp Ala Leu Arg Glu Ser Glu Phe Thr Arg Trp Val Tyr
 515 520 525
 Gln Gly Leu Cys Gly Leu Glu Lys Glu Gln Lys Asp Ala Leu
 530 535 540

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 250 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

- (ii) MOLECULE TYPE: protein

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Gly Leu Thr Thr Lys Pro Leu Ser Leu Lys Val Asn Ala Ala Leu
 1 5 10 15
 Phe Asp Val Asp Gly Thr Ile Ile Ile Ser Gln Pro Ala Ile Ala Ala
 20 25 30
 Phe Trp Arg Asp Phe Gly Lys Asp Lys Pro Tyr Phe Asp Ala Glu His
 35 40 45
 Val Ile Gln Val Ser His Gly Trp Arg Thr Phe Asp Ala Ile Ala Lys
 50 55 60
 Phe Ala Pro Asp Phe Ala Asn Glu Glu Tyr Val Asn Lys Leu Glu Ala
 65 70 75 80
 Glu Ile Pro Val Lys Tyr Gly Glu Lys Ser Ile Glu Val Pro Gly Ala
 85 90 95

GAGCAGCTT GCGGAGGCGG

Val Lys Leu Cys Asn Ala Leu Asn Ala Leu Pro Lys Glu Lys Trp Ala
100 105 110
Val Ala Thr Ser Gly Thr Arg Asp Met Ala Gln Lys Trp Phe Glu His
115 120 125
Leu Gly Ile Arg Arg Pro Lys Tyr Phe Ile Thr Ala Asn Asp Val Lys
130 135 140
Gln Gly Lys Pro His Pro Glu Pro Tyr Leu Lys Gly Arg Asn Gly Leu
145 150 155 160
Gly Tyr Pro Ile Asn Glu Gln Asp Pro Ser Lys Ser Lys Val Val Val
165 170 175
Phe Glu Asp Ala Pro Ala Gly Ile Ala Ala Gly Lys Ala Ala Gly Cys
180 185 190
Lys Ile Ile Gly Ile Ala Thr Thr Phe Asp Leu Asp Phe Leu Lys Glu
195 200 205
Lys Gly Cys Asp Ile Ile Val Lys Asn His Glu Ser Ile Arg Val Gly
210 215 220
Gly Tyr Asn Ala Glu Thr Asp Glu Val Glu Phe Ile Phe Asp Asp Tyr
225 230 235 240
Leu Tyr Ala Lys Asp Asp Leu Leu Lys Trp
245 250

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 271 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Lys Arg Phe Asn Val Leu Lys Tyr Ile Arg Thr Thr Lys Ala Asn
1 5 10 15
Ile Gln Thr Ile Ala Met Pro Leu Thr Thr Lys Pro Leu Ser Leu Lys
20 25 30
Ile Asn Ala Ala Leu Phe Asp Val Asp Gly Thr Ile Ile Ser Gln
35 40 45
Pro Ala Ile Ala Ala Phe Trp Arg Asp Phe Gly Lys Asp Lys Pro Tyr
50 55 60
Phe Asp Ala Glu His Val Ile His Ile Ser His Gly Trp Arg Thr Tyr
65 70 75 80

Asp Ala Ile Ala Lys Phe Ala Pro Asp Phe Ala Asp Glu Glu Tyr Val
 85 90 95
 Asn Lys Leu Glu Gly Glu Ile Pro Glu Lys Tyr Gly Glu His Ser Ile
 100 105 110
 Glu Val Pro Gly Ala Val Lys Leu Cys Asn Ala Leu Asn Ala Leu Pro
 115 120 125
 Lys Glu Lys Trp Ala Val Ala Thr Ser Gly Thr Arg Asp Met Ala Lys
 130 135 140
 Lys Trp Phe Asp Ile Leu Lys Ile Lys Arg Pro Glu Tyr Phe Ile Thr
 145 150 155 160
 Ala Asn Asp Val Lys Gln Gly Lys Pro His Pro Glu Pro Tyr Leu Lys
 165 170 175
 Gly Arg Asn Gly Leu Gly Phe Pro Ile Asn Glu Gln Asp Pro Ser Lys
 180 185 190
 Ser Lys Val Val Val Phe Glu Asp Ala Pro Ala Gly Ile Ala Ala Gly
 195 200 205
 Lys Ala Ala Gly Cys Lys Ile Val Gly Ile Ala Thr Thr Phe Asp Leu
 210 215 220
 Asp Phe Leu Lys Glu Lys Gly Cys Asp Ile Ile Val Lys Asn His Glu
 225 230 235 240
 Ser Ile Arg Val Gly Glu Tyr Asn Ala Glu Thr Asp Glu Val Glu Leu
 245 250 255
 Ile Phe Asp Asp Tyr Leu Tyr Ala Lys Asp Asp Leu Leu Lys Trp
 260 265 270

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 709 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met Phe Pro Ser Leu Phe Arg Leu Val Val Phe Ser Lys Arg Tyr Ile
 1 5 10 15
 Phe Arg Ser Ser Gln Arg Leu Tyr Thr Ser Leu Lys Gln Glu Gln Ser
 20 25 30
 Arg Met Ser Lys Ile Met Glu Asp Leu Arg Ser Asp Tyr Val Pro Leu
 35 40 45

Ile Ala Ser Ile Asp Val Gly Thr Thr Ser Ser Arg Cys Ile Leu Phe
50 55 60

Asn Arg Trp Gly Gln Asp Val Ser Lys His Gln Ile Glu Tyr Ser Thr
65 70 75 80

Ser Ala Ser Lys Gly Lys Ile Gly Val Ser Gly Leu Arg Arg Pro Ser
85 90 95

Thr Ala Pro Ala Arg Glu Thr Pro Asn Ala Gly Asp Ile Lys Thr Ser
100 105 110

Gly Lys Pro Ile Phe Ser Ala Glu Gly Tyr Ala Ile Gln Glu Thr Lys
115 120 125

Phe Leu Lys Ile Glu Glu Leu Asp Leu Asp Phe His Asn Glu Pro Thr
130 135 140

Leu Lys Phe Pro Lys Pro Gly Trp Val Glu Cys His Pro Gln Lys Leu
145 150 155 160

Leu Val Asn Val Val Gln Cys Leu Ala Ser Ser Leu Leu Ser Leu Gln
165 170 175

Thr Ile Asn Ser Glu Arg Val Ala Asn Gly Leu Pro Pro Tyr Lys Val
180 185 190

Ile Cys Met Gly Ile Ala Asn Met Arg Glu Thr Thr Ile Leu Trp Ser
195 200 205

Arg Arg Thr Gly Lys Pro Ile Val Asn Tyr Gly Ile Val Trp Asn Asp
210 215 220

Thr Arg Thr Ile Lys Ile Val Arg Asp Lys Trp Gln Asn Thr Ser Val
225 230 235 240

Asp Arg Gln Leu Gln Leu Arg Gln Lys Thr Gly Leu Pro Leu Leu Ser
245 250 255

Thr Tyr Phe Ser Cys Ser Lys Leu Arg Trp Phe Leu Asp Asn Glu Pro
260 265 270

Leu Cys Thr Lys Ala Tyr Glu Glu Asn Asp Leu Met Phe Gly Thr Val
275 280 285

Asp Thr Trp Leu Ile Tyr Gln Leu Thr Lys Gln Lys Ala Phe Val Ser
290 295 300

Asp Val Thr Asn Ala Ser Arg Thr Gly Phe Met Asn Leu Ser Thr Leu
305 310 315 320

Lys Tyr Asp Asn Glu Leu Leu Glu Phe Trp Gly Ile Asp Lys Asn Leu
325 330 335

Ile His Met Pro Glu Ile Val Ser Ser Ser Gln Tyr Tyr Gly Asp Phe
340 345 350

355

360

365

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Val | Leu | | Asp | Leu | Val | Lys | Arg | Asn | Leu | Phe | Ile | Gln | Gly | Cys | Leu |
| 370 | | | | | | 375 | | | | | | 380 | | | |
| Gly | Asp | Gln | Ser | Ala | Ser | Met | Val | Gly | Gln | Leu | Ala | Tyr | Lys | Pro | Gly |
| 385 | | | | | 390 | | | | | 395 | | | | | 400 |
| Ala | Ala | Lys | Cys | Thr | Tyr | Gly | Thr | Gly | Cys | Phe | Leu | Leu | Tyr | Asn | Thr |
| | | | | 405 | | | | | 410 | | | | | 415 | |
| Gly | Thr | Lys | Lys | Leu | Ile | Ser | Gln | His | Gly | Ala | Leu | Thr | Thr | Leu | Ala |
| | | | | 420 | | | | 425 | | | | | | 430 | |
| Phe | Trp | Phe | Pro | His | Leu | Gln | Glu | Tyr | Gly | Gly | Gln | Lys | Pro | Glu | Leu |
| | | | | 435 | | | 440 | | | | 445 | | | | |
| Ser | Lys | Pro | His | Phe | Ala | Leu | Glu | Gly | Ser | Val | Ala | Val | Ala | Gly | Ala |
| | 450 | | | | | 455 | | | | 460 | | | | | |
| Val | Val | Gln | Trp | Leu | Arg | Asp | Asn | Leu | Arg | Leu | Ile | Asp | Lys | Ser | Glu |
| | 465 | | | | | 470 | | | | 475 | | | | | 480 |
| Asp | Val | Gly | Pro | Ile | Ala | Ser | Thr | Val | Pro | Asp | Ser | Gly | Gly | Val | Val |
| | | | | 485 | | | | 490 | | | | | | 495 | |
| Phe | Val | Pro | Ala | Phe | Ser | Gly | Leu | Phe | Ala | Pro | Tyr | Trp | Asp | Pro | Asp |
| | | | | 500 | | | | 505 | | | | 510 | | | |
| Ala | Arg | Ala | Thr | Ile | Met | Gly | Met | Ser | Gln | Phe | Thr | Thr | Ala | Ser | His |
| | | | | 515 | | | | 520 | | | | 525 | | | |
| Ile | Ala | Arg | Ala | Ala | Val | Glu | Gly | Val | Cys | Phe | Gln | Ala | Arg | Ala | Ile |
| | | | | 530 | | | 535 | | | | 540 | | | | |
| Leu | Lys | Ala | Met | Ser | Ser | Asp | Ala | Phe | Gly | Glu | Gly | Ser | Lys | Asp | Arg |
| | 545 | | | | | 550 | | | | 555 | | | | 560 | |
| Asp | Phe | Leu | Glu | Glu | Ile | Ser | Asp | Val | Thr | Tyr | Glu | Lys | Ser | Pro | Leu |
| | | | | | 565 | | | 570 | | | | | | 575 | |
| Ser | Val | Leu | Ala | Val | Asp | Gly | Gly | Met | Ser | Arg | Ser | Asn | Glu | Val | Met |
| | | | | 580 | | | | 585 | | | | | 590 | | |
| Gln | Ile | Gln | Ala | Asp | Ile | Leu | Gly | Pro | Cys | Val | Lys | Val | Arg | Arg | Ser |
| | | | | 595 | | | | 600 | | | | 605 | | | |
| Pro | Thr | Ala | Glu | Cys | Thr | Ala | Leu | Gly | Ala | Ala | Ile | Ala | Ala | Asn | Met |
| | | | | 610 | | | 615 | | | | 620 | | | | |
| Ala | Phe | Lys | Asp | Val | Asn | Glu | Arg | Pro | Leu | Trp | Lys | Asp | Leu | His | Asp |
| | | | | 625 | | | 630 | | | 635 | | | | 640 | |
| Val | Lys | Lys | Trp | Val | Phe | Tyr | Asn | Gly | Met | Glu | Lys | Asn | Glu | Gln | Ile |
| | | | | | 645 | | | | 650 | | | | 655 | | |

Ser Pro Glu Ala His Pro Asn Leu Lys Ile Phe Arg Ser Glu Ser Asp
660 665 670

Asp Ala Glu Arg Arg Lys His Trp Lys Tyr Trp Glu Val Ala Val Glu
675 680 685

Arg Ser Lys Gly Trp Leu Lys Asp Ile Glu Gly Glu His Glu Gln Val
690 695 700

Leu Glu Asn Phe Gln
705

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GCGCGGATCC AGGAGTCTAG AATTATGGGA TTGACTACTA AACCTCTATC T

51

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GATACGCCCG GGTTACCATT TCAACAGATC GTCCTT

36

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TTGATAATAT ACCCATGGCT GCTGCTGCTG ATAG

34

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GTATGATATG TTATCTTGGAA TCCAATAAAAT CTAATCTTC

39

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CATGACTAGT AAGGAGGACA ATTC

24

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CATGGAATTG TCCTCCTTAC TAGT

24

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CTAGTAAGGA GGACAATTG

19

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CATGGAATTG TCCTCCTTA

19

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GATCCAGGAA ACAGA

15

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CTAGTCTGTT TCCTG

15

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GCTTTCTGTG CTGCGGCTTT AG

22

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TGGTCGAGGA TCCACTTCAC TTT

23

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

AAAGTGAAGT GGATCCTCGA CCAATTGGAT GGTGGCGCAG TAGCAAACAA T

51

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GGATCACCGC CGCAGAACT ACG

23

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CTGTCAGCCG TTAAGTGTTC CTGTG

25

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CAGTTCAACC TGTTGATAGT ACG

23

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ATGAGTCAAA CATAACACCTT

20

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

ATGGAGAAAA AAATCACTGG

20

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

TTACGCCCG CCCTGCCACT

20

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TCAGAGGATG TGCACCTGCA

20

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

CGAGCATGCC GCATTTGGCA CTACTC

26

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GCGTCTAGAG TAGGTTATTG CCACTCTTG

29

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GAAGTCGACC GCTGCGCCTT ATCCGG

26

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs

GCAGCATGCT GGACTGGTAG TAG

28

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:
CGCGTCGACG TTTACAATTT CAGGTGGC

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:
GCAGCATGCT GGACTGGTAG TAG

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:
CAGTCTAGAG TTATTGGCAA ACCTACC

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:
GATGCATGCC CAGGGCGGAG ACGGC

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

CTAACGATTG TTCTCTAGAG AAAATGTCC

29